

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 06-282293

(43)Date of publication of application : 07.10.1994

(51)Int.Cl. G10L 3/00
G10L 5/06

(21)Application number : 05-069565

(71)Applicant : SONY CORP

(22)Date of filing : 29.03.1993

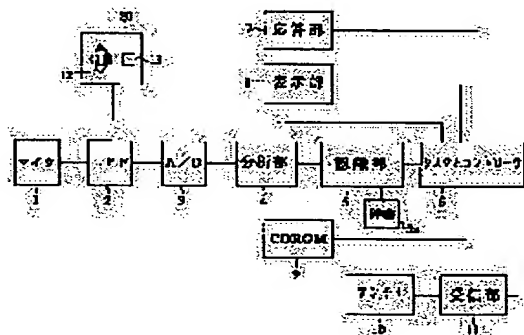
(72)Inventor : OMOTE MASANORI

(54) VOICE RECOGNITION DEVICE

(57)Abstract:

PURPOSE: To select a correct recognized result easily from recognized result candidates by classifying and displaying the outputted recognizing result candidates according to a prescribed constitutive unit after a voice is recognized according to an analysis result of an inputted voice.

CONSTITUTION: For example, in a navigation system, a recognizing part 5 recognizes continuously a voice inputted to a microphone 1 from an analysis result of an analytical part 4 according to a standard pattern and a sentence structure registered beforehand in a dictionary 5a, and outputs the recognized result candidates to a system controller 6. The system controller 6 classifies these recognized result candidates with every prescribed constitutive unit such as, for example, the character number, and displays the recognized result candidate of a character designated (selected) by a joy stick 12 on a display part 8, for example, in order of high tolerance. The system controller 6 controls the whole device according to a control signal outputted from an operation part 20 by operation such as the joy stick 12 and a recognized result determined from the recognized result candidates of the recognizing part 5.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's
decision of rejection]

[Date of extinction of right]

(11)特許出願公開番号

特開平6-282293

(43)公開日 平成6年(1994)10月7日

(51)Int.Cl.⁵

G 1 0 L 3/00

識別記号

561 C

庁内整理番号

9379-5H

E 9379-5H

E 9379-5H

FI

技術表示箇所

審査請求 未請求 請求項の数3 OL (全 9 頁)

(21)出願番号

特願平5-69565

(22)出題日

平成5年(1993)3月29日

(71)出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72)発明者 表 雅則

東京都品川区北品川6丁目7番35号 ソニ

一株式会社内

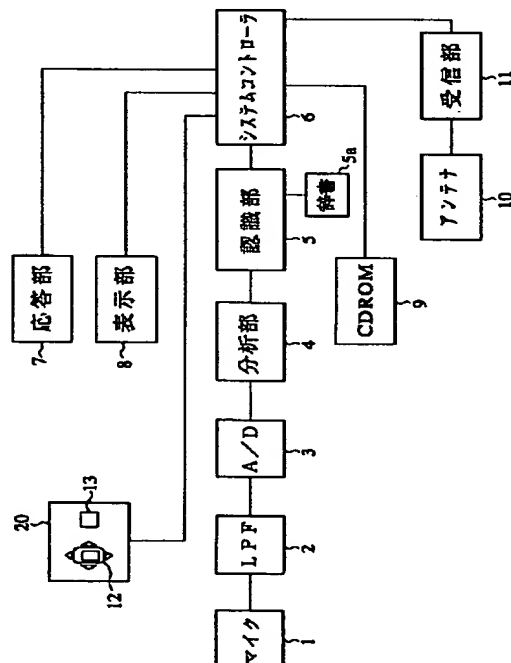
(74)代理人 弁理士 稻本 義雄

(54)【発明の名称】 音声認識装置

(57) 【要約】

【目的】 認識結果候補の中から、正しい認識結果を容易に選択することができるようにする。

【構成】 マイク1、LPF2、A/D変換器3を介して入力された音声が入力部4で分析され、認識部5において、その分析結果に基づいて、音声が入力部5で認識され、認識結果候補が出力される。そして、システムコントローラ6において、その認識結果候補が、その文字数に基づいて分類され、尤度の最も大きい認識結果候補と同一の文字数の認識結果候補が、その尤度の大きい順に、表示部8に表示される。



【特許請求の範囲】

【請求項1】 入力された音声进行分析する分析手段と、前記分析手段の分析結果に基づいて、前記音声認識し、その認識結果候補を出力する認識手段と、前記認識手段より出力された前記認識結果候補を、その所定の構成単位数に基づいて分類する分類手段と、前記分類手段の分類結果を表示する表示手段とを備えることを特徴とする音声認識装置。

【請求項2】 前記所定の構成単位数を選択する選択手段をさらに備え、

前記表示手段は、前記選択手段により選択された前記所定の構成単位数の前記認識結果候補を表示することを特徴とする請求項1に記載の音声認識装置。

【請求項3】 前記表示手段は、前記分類手段により分類された前記認識結果候補のうちの最も尤度の大きいものと同一の所定の構成単位数の前記認識結果候補を表示することを特徴とする請求項1に記載の音声認識装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、音声認識する場合に用いて好適な音声認識装置に関する。

【0002】

【従来の技術】 図8は、従来の音声認識装置の一例の構成を示すブロック図である。マイク1は、入力された音声信号を電気信号としての音声信号に変換する。LPF（ローパスフィルタ）2は、入力された音声信号の周波数帯域を、次段のA/D変換器3におけるサンプリング周波数のほぼ1/2以下に制限する。A/D変換器3は、入力された音声信号を所定のサンプリング周波数でサンプリングし、分析部4に出力する。分析部4は、入力された音声信号を分析し、例えば線形予測係数やケプストラム係数、所定の周波数帯域ごとのパワー、ゼロクロス数などの音声の特徴量を抽出する。認識部31は、辞書31aにあらかじめ記憶された標準パターンと、分析部4から出力される音声の特徴量としての入力パターンとを照合し、入力パターンと類似する（同一である）尤度が、例えば最も大きい標準パターンを音声認識結果として出力する。

【0003】 以上のように構成される音声認識装置では、マイク1に音声が入力されると、その音声は、LPF2で帯域制限され、A/D変換器3でA/D変換される。A/D変換された音声信号は、分析部4に入力され、そこで音声の特徴パラメータが抽出されて、認識部31に出力される。認識部31では、辞書31aが参照され、辞書31aにあらかじめ記憶された標準パターンと、分析部4から出力される音声の特徴量としての入力パターンとが照合される。そして、入力パターンと類似する尤度が最も大きい標準パターンが認識結果として出力される。

【0004】 ところで、このような音声認識装置におい

ては、認識の対象とする単語の数が増加すると、即ち辞書31aにあらかじめ登録しておく標準パターンの数が増加すると、入力された音声に対して大きい尤度を有する標準パターン間において、尤度の差がほとんどなくなり、尤度の最も大きい標準パターンが認識結果として正しいとは限らない場合があった。

【0005】 即ち、例えば雑音を含む音声や、発声の曖昧な音声などが入力された場合、認識結果として正しい標準パターンの尤度が、他の標準パターンの尤度より小さくなり、その結果、誤認識される問題があった。

【0006】 そこで、尤度が大きい順に、所定の数だけの標準パターンが認識結果候補として出力されるように認識部31を構成し、この所定の数だけの標準パターンを、ディスプレイなどに表示して、使用者に正しいものを選択させる方法がある。

【0007】

【発明が解決しようとする課題】 しかしながら、この方法では、辞書31aに登録しておく標準パターンを増加し、この結果、類似した標準パターンが多くなり、誤認識が生じ易くなると、認識結果候補として表示する標準パターンの数を多くしなければならなかった。従って、この場合、使用者は、多くの認識結果候補の中から、正しいものを探して選択しなければならず、使用者にわずらわしさを感じさせる課題があった。

【0008】 以上は、入力された音声から1単語（または音韻や文節など）を認識する離散単語認識の場合であるが、連続音声認識においても、上述の単語認識の技術が適用されるので、やはり同様の課題があった。

【0009】 さらに、単語認識の技術が適用された連続音声認識においては、入力音声の始まりの部分から、標準パターンを入力音声の一部としての部分パターンとして照合することを、入力音声の終わりの部分まで繰り返すようになされている。

【0010】 即ち、連続音声認識では、例えば標準パターンとしてA、B、C、D、Eが用意されていた場合、入力音声の始まりの部分に、標準パターンA、B、C、D、Eが、その部分パターンとして照合され、その結果、尤度の比較的大きい標準パターン（例えば、標準パターンA、B、C）が、入力音声の始まりの部分の認識結果候補とされる。

【0011】 さらに、入力音声の続き（いまの場合、入力音声から、その始まりの部分の認識結果候補としての標準パターンA、B、またはCに対応する部分を除いた残り）の部分に、標準パターンA、B、C、D、Eが、その部分パターンとして、それぞれ照合され、その結果、尤度の比較的大きい標準パターンが、入力音声の続きの部分の認識結果候補とされる。

【0012】 以下同様にして、入力音声の終わりの部分まで、入力音声の続きに標準パターンを、その部分パターンとして照合することが繰り返される。

【0013】従って、連続音声認識においては、尤度の大きい所定数（例えば、上位3つ）の標準パターンを認識結果候補とするようにした場合、入力音声の始まりの部分の認識結果候補として、例えば標準パターンA、B、Cが出力され、さらにこれらの標準パターンA、B、Cそれぞれに続く認識結果候補として、尤度の大きい3つの標準パターンが出力されることになる。

【0014】即ち、この場合、入力音声の認識結果候補の数は、3のべき乗で増加することになる。

【0015】例えばATR（国際電気通信基礎技術研究所）で開発されたHMM—LR方式（HMM（Hidden Markov Model）法による音声認識と、LRテーブルによる構文解析を組み合わせた手法）などによる連続音声認識では、以上のように、尤度の大きい所定数の標準パターンを認識結果候補として記憶しておき、それに続く（入力音声の）部分と標準パターンの照合を、順次行うようになされている。

【0016】従って、認識結果候補として、多くの標準パターンを記憶しておかなければならず、装置が大型化する課題があった。さらに、多くの認識結果候補の組み合わせの中から、正確な認識結果を選択しなければならず、正確な認識結果を得るのに時間がかかる課題があった。

【0017】本発明は、このような状況に鑑みてなされたものであり、正確な認識結果を容易に、且つ短時間で得ることができるようにするものである。

【0018】

【課題を解決するための手段】請求項1に記載の音声認識装置は、入力された音声进行分析する分析手段としての分析部4と、分析部4の分析結果に基づいて、音声を認識し、その認識結果候補を出力する認識手段としての認識部5と、認識部5より出力された認識結果候補を、例えば文字数などの、その所定の構成単位数に基づいて分類する分類手段としてのシステムコントローラ6と、システムコントローラ6の分類結果を表示する表示手段としての表示部8とを備えることを特徴とする。

【0019】請求項2に記載の音声認識装置は、所定の構成単位数を選択する選択手段としてのジョイスティック12をさらに備え、表示部8に、ジョイスティック12により選択された所定の構成単位数の認識結果候補を表示させることを特徴とする。

【0020】請求項3に記載の音声認識装置は、表示部8に、システムコントローラ6により分類された認識結果候補のうちの最も尤度の大きいものと同一の所定の構成単位数の認識結果候補を表示させることを特徴とする。

【0021】

【作用】請求項1に記載の音声認識装置においては、入力された音声分析され、その分析結果に基づいて、音声認識されて、認識結果候補が出力される。そして、

その認識結果候補が、例えば文字数などの所定の構成単位数に基づいて分類されて表示される。従って、認識結果候補の中から、正しい認識結果を容易に選択することができ、正しい認識結果を短時間で得ることができる。

【0022】請求項2に記載の音声認識装置においては、表示部8に、ジョイスティック12により選択された所定の構成単位数の認識結果候補を表示させるので、認識結果候補の中から、正しい認識結果を、さらに容易に選択することができる。

【0023】請求項3に記載の音声認識装置においては、表示部8に、システムコントローラ6により分類された認識結果候補のうちの最も尤度の大きいものと同一の所定の構成単位数の認識結果候補を表示させる。従って、より確からしい認識結果候補が表示されるので、正しい認識結果を、さらに短時間で得ることができる。

【0024】

【実施例】図1は、本発明の音声認識装置を適用したナビゲーションシステムの一実施例の構成を示すブロック図である。図中、図8の音声認識装置における場合と対応する部分については、同一の符号を付してある。

【0025】認識部5は、辞書5aにあらかじめ登録された標準パターンおよび構文情報に基づいて、分析部4の分析結果から、マイク1に入力された音声を連続音声認識し、その認識結果候補をシステムコントローラ6に出力する。辞書5aには、標準パターンおよび構文情報があらかじめ記憶されている。

【0026】システムコントローラ6は、認識部5からの認識結果候補を、例えばその文字数などの所定の構成単位数ごとに分類し、ジョイスティック12によって指定（選択）される文字数の認識結果候補を、例えば尤度の高い順に表示部8に表示させる。

【0027】さらに、システムコントローラ6は、ジョイスティック12もしくはボタン13などが操作されることにより操作部20から出力される制御信号、または認識部5の認識結果候補から決定された認識結果に基づいて、装置全体を制御する。応答部7は、システムコントローラ6からの指示にしたがって、例えば道案内のための音声などの所定の音声出力する（音声応答する）。表示部8は、システムコントローラ6からの指示にしたがって、CDROM9に記憶された地図情報など、所定の情報を表示する。さらに、表示部8は、システムコントローラ6からの、所定の文字数ごとの認識結果候補を表示する。CDROM9には、地図情報や交通情報などがあらかじめ記憶されている。アンテナ10は、GPS（図示せず）からの電波を受信し、受信部11に出力する。受信部11は、アンテナ10からの信号を復調する。

【0028】操作部20のジョイスティック12は、マイク1からの音声入力が無効にされている場合（装置が、マイク1からの音声を認識するモードになっていな

い場合)、表示部8に表示された地図を左、右、上、下にスクロールするときや、表示部8のカーソルを左、右、上、下に移動するときなどに、左、右、上、下にそれぞれ操作される。

【0029】さらに、この場合、ジョイスティック12は、例えば表示部8に表示された地図上の目的地を設定するときや、表示部8に表示されたメニューの中の項目を選択後、確定するときなどに押し込まれるように操作される(押し込まれる)。

【0030】また、ジョイスティック12は、マイク1からの音声入力が無効にされている場合、(装置が、マイク1からの音声を認識するモードになっている場合)、表示部8に表示させる認識結果候補の文字数を増加または減少させるときや、所定の入力音声の位置における部分パターンとしての認識結果候補を選択した後、次に続く入力音声の位置における部分パターンとしての認識結果候補を表示部8に表示させるか、または所定の入力音声の位置における認識結果候補の選択を中止し、その直前の入力音声の位置における認識結果候補を選択する状態にするときなどに、右または左にそれぞれ操作される。

【0031】さらに、この場合、ジョイスティック12は、表示部8に表示された認識結果候補を選択するために、カーソルを上下するとき、上下に操作され、選択した認識結果候補が正しく、それを仮に確定するときには、押し込まれる。

【0032】操作部20のボタン13は、マイク1からの音声入力を無効/有効にするとときに操作される。

【0033】なお、このナビゲーションシステムには、以上のようなジョイスティック12およびボタン13の他、例えば表示部13に表示された地図を拡大または縮小したりするなど、所定の動作を行わせるための、図示せぬボタン類が設けられている。

【0034】以上のように構成されるナビゲーションシステムは、自動車などに搭載され、操作部20のジョイスティック12またはボタン13の操作だけでなく、マイク1に入力される音声により所定の動作(例えば、地図や現在地の表示、設定された目的地までの誘導(道案内)など)をさせることができるようになっている。

【0035】即ち、まず、マイク1からの音声入力が無効にされていない場合、システムコントローラ6において、ジョイスティック12や他の図示せぬボタン類の操作に対応して、CDROM9に記憶された地図情報が読み出され、表示部8に表示される。また、アンテナ10および受信部11を介して入力されるGPSからの信号を基に、現在地が算出され、表示部8に表示される。あるいは、例えば地図情報表示や、現在地表示、目的地の設定、目的地までの道案内などのメニューの選択、表示された地図情報の拡大/縮小/スクロールなどが行われる。

【0036】次に、ボタン13が操作され、マイク1からの音声入力が無効にされた場合の動作について、図2のフローチャートを参照して説明する。この場合、まずステップS1において、マイク1に音声が入力されると、LPF2およびA/D変換器3を介して分析部4に出力されて分析される。分析部4での分析結果は、認識部5に入力される。

【0037】そして、ステップS2において、認識部5で、辞書5aが参照され、入力音声の始まりの部分に、辞書5aに記憶された標準パターンすべてが、その部分パターンとして、順次照合され、その結果、尤度の大きい上位N個の標準パターンが、入力音声の始まりの部分の認識結果候補とされる。

【0038】この尤度の大きい上位N個の標準パターン(即ち、認識結果候補)は、システムコントローラ6に入力され、ステップS3において、システムコントローラ6で、認識部5からの認識結果候補のうち、尤度の最も大きい認識結果候補と同一の文字数Cの認識結果候補が検出され、ステップS4に進む。

【0039】ステップS4において、ステップS3で検出された尤度の最も大きい認識結果候補と同一の文字数Cの認識結果候補(認識結果候補に対応する単語)が、その尤度の大きい順に表示部8に表示され、ステップS5に進み、操作部20のジョイスティック12が操作され、システムコントローラ6に制御信号の入力があつたか否かが判定される。

【0040】ステップS5において、ジョイスティック12が操作されておらず、システムコントローラ6に制御信号の入力がなかったと判定された場合、再びステップS5に戻る。ステップS5において、ジョイスティック12が操作され、システムコントローラ6に制御信号が入力されたと判定された場合、ステップS6に進み、その制御信号が、ジョイスティック12が右に操作されたことによるものであるか否かが判定される。

【0041】ステップS6において、ジョイスティック12が右に操作されたことによる制御信号が、システムコントローラ6に入力されたと判定された場合、ステップS7に進み、いま表示している認識結果候補の文字数Cが1だけインクリメントされ、認識部5からの認識結果候補のうち、文字数がC+1であるものが検出され、ステップS4に戻る。

【0042】従って、この場合、ステップS4において、文字数がC+1の認識結果候補が、文字数がCの認識結果候補に代わって、表示部8に表示されることになる。

【0043】一方、ステップS6において、システムコントローラ6に入力された制御信号が、ジョイスティック12が右に操作されたことによるものでないと判定された場合、ステップS8に進み、その制御信号が、ジョイスティック12が左に操作されたことによるものであ

るか否かが判定される。

【0044】ステップS8において、ジョイスティック12が左に操作されたことによる制御信号が、システムコントローラ6に入力されたと判定された場合、ステップS9に進み、いま表示している認識結果候補の文字数Cが1だけデクリメントされ、認識部5からの認識結果候補のうち、文字数がC－1であるものが検出され、ステップS4に戻る。

【0045】従って、この場合、ステップS4において、文字数がC－1の認識結果候補が、文字数がCの認識結果候補に代わって、表示部8に表示されることになる。

【0046】一方、ステップS8において、システムコントローラ6に入力された制御信号が、ジョイスティック12が左に操作されたことによるものでないと判定された場合、ステップS10に進み、その制御信号が、ジョイスティック12が上に操作されたことによるものであるか否かが判定される。

【0047】ステップS10において、ジョイスティック12が上に操作されたことによる制御信号が、システムコントローラ6に入力されたと判定された場合、ステップS11に進み、表示部8に表示されているカーソルが、1行上の認識結果候補に移動され、ステップS4に戻る。

【0048】なお、この場合、表示部8の画面の最上行にカーソルが表示されており、且つその最上行に表示されている認識結果候補より尤度の大きい認識結果候補がある場合には、その尤度の大きい認識結果候補が表示されるように、画面が下に1行だけスクロールするようになされている。

【0049】一方、ステップS10において、システムコントローラ6に入力された制御信号が、ジョイスティック12が上に操作されたことによるものでないと判定された場合、ステップS12に進み、その制御信号が、ジョイスティック12が下に操作されたことによるものであるか否かが判定される。

【0050】ステップS12において、ジョイスティック12が下に操作されたことによる制御信号が、システムコントローラ6に入力されたと判定された場合、ステップS13に進み、表示部8に表示されているカーソルが、1行下の認識結果候補に移動され、ステップS4に戻る。

【0051】なお、この場合、表示部8の画面の最下行にカーソルが表示されており、且つその最下行に表示されている認識結果候補より尤度の小さい認識結果候補がある場合には、その尤度の小さい認識結果候補が表示されるように、画面が上に1行だけスクロールするようになされている。

【0052】一方、ステップS12において、システムコントローラ6に入力された制御信号が、ジョイスティ

ック12が下に操作されたことによるものでないと判定された場合、ステップS14に進み、その制御信号が、ジョイスティック12が押し込まれたことによるものであるか否かが判定される。

【0053】ステップS14において、ジョイスティック12が押し込まれたことによる制御信号が、システムコントローラ6に入力されたと判定された場合、ステップS15に進み、いまカーソルのある位置に表示されている認識結果候補としての標準パターン（標準パターンに対応する単語）が、入力音声の始まりの部分の認識結果として仮に決定（確定）され、この仮決定（仮確定）された標準パターンに続く入力音声の部分に注目点が移行されて、ステップS2に戻る。

【0054】そして、ステップS2以下では、上述した場合と同様にして、入力音声の終わりの部分までに対応する部分パターンとしての認識結果候補が順次仮決定されることになる。

【0055】一方、ステップS14において、システムコントローラ6に入力された制御信号が、ジョイスティック12が押し込まれたことによるものでないと判定された場合、ステップS16に進み、所定のエラー処理が行われ、ステップS1に戻る。

【0056】以上のようにして、入力音声に対応する標準パターンの仮決定がなされた後、その仮決定結果が入力音声の認識結果として正しければ、図示せぬ確定ボタンが操作され、仮決定結果が確定される。そして、その確定結果、即ち入力音声の正しい認識結果に基づいて、システムコントローラ6により装置が制御される。

【0057】即ち、認識結果が、例えば地名であった場合、システムコントローラ6において、その地名付近の地図が表示部8に表示されたり、あるいはその地名が目的地として設定される。また、認識結果が、例えばコマンドであった場合、そのコマンドに対応して、地図情報や現在地情報が表示部8に表示されたり、メニューの選択、表示された地図情報の拡大／縮小／スクロールなどが行われる。

【0058】なお、音声によるコマンド入力があった場合、必要に応じて、音声応答部7からは、音声による応答（例えば入力された音声の確認や、さらなるコマンドの入力の催促）が出力されるようになされており、これにより使用者が、対話形式で装置を使用することができるようになされている。

【0059】次に、図3乃至図7を参照して、マイク1からの音声入力が有効にされた場合の動作について、さらに説明する。ボタン13が操作され、マイク1からの音声入力が有効にされると、まず、表示部8に、例えば図3に示すような、その下部にインジケータ22および候補確定窓23を有する音声入力窓21が表示される。

【0060】そして、所定の地域の地図を表示させるために、例えばマイク1に「東京都品川区北品川」などの

連続音声が入力されると、この入力音声は、LPF2、A/D変換器3、および分析部4を介して認識部5に出力される。認識部5では、その内蔵する音声区間検出回路(図示せず)によって、マイク1に入力された連続音声「東京都品川区北品川」の音声区間(発話時間)が検出され、システムコントローラ6に出力される。システムコントローラ6は、例えば図4に示すように、認識部5からの、入力音声「東京都品川区北品川」の音声区間(発話時間)に対応して、表示部8のインジケータ22を伸ばして表示させる。

【0061】同時に、入力音声「東京都品川区北品川」は、LPF2およびA/D変換器3を介して分析部4に出力されて分析される。分析部4での分析結果は、認識部5に入力される。

【0062】そして、認識部5で、辞書5aの標準パターンおよび構文情報が参照され、入力音声の始まりの部分に、辞書5aに記憶された標準パターンすべてが、その部分パターンとして、順次照合され、その結果、尤度の大きい上位N個の標準パターンが、入力音声の始まりの部分の認識結果候補とされる。

【0063】この尤度の大きい上位N個の標準パターン(即ち、認識結果候補)は、システムコントローラ6に入力され、そのうち、尤度の最も大きい認識結果候補と同一の文字数Cの認識結果候補が検出される。そして、図4に示すように、この文字数Cの認識結果候補(認識結果候補に対応する単語)が、その尤度の大きい順に表示部8の音声入力窓21内にオープンされたウインドウ24aに表示される。

【0064】即ち、認識結果候補が、その尤度の大きい順に、例えば「東京都」、「東京タワー」、「神奈川県」、「和歌山県」、「東京駅」、「京都府」、……であったとした場合、従来の装置では、この通りに表示されるが、この音声認識装置では、図4に示すように、入力音声の始まりの部分の認識結果候補のうち、尤度が最も高かった認識結果候補「東京都」がウインドウ24aの最上行に表示され、2行目以降には、2乃至4番目に尤度の高い、5文字の「東京タワー」、4文字の「神奈川県」および「和歌山県」は表示されずに、それより尤度の低い、認識結果候補「東京都」と同一の文字数の認識結果候補(図4においては、「東京駅」、「京都府」、……)が尤度の大きい順に表示される。

【0065】さらに、ウインドウ24aの最上行(この場合には、「東京都」が表示されている行)にはカーソル(図中、影を付してある部分)が表示され、このカーソルが位置している認識結果候補が、候補確定窓23に表示される。

【0066】ところで、図4においては、入力音声「東京都品川区北品川」の始まりの部分の認識結果として、3文字の認識結果候補が表示されているが、例えば入力音声の始まりの部分の認識結果として、例えば4文字の

単語が正しい認識結果である場合に、使用者によって、ジョイスティック12が右に操作されると、図5に示すように、4文字の認識結果候補としての「神奈川県」、「和歌山県」、……が尤度の大きい順に表示される。

【0067】従って、使用者は、自身が発した音声であるから、その正しい認識結果の文字数を、ジョイスティック12により容易に指定することができ、ウインドウ24aには、指定された文字数の認識結果候補のみが表示されるので、従来のように、尤度の高い順に所定数の認識結果候補が表示される場合に比較して、正しい認識結果を、容易に且つ短時間に選択して得ることができるようになる。

【0068】なお、上述したように、図4に示す状態において、ジョイスティック12が左に操作されると、2文字の認識結果候補が、尤度の大きい順に表示される。また、ジョイスティック12により指定された文字数以上、または以下の認識結果候補が存在しない場合には、その直前に表示されていた文字数の認識結果候補がそのまま表示され続ける。

【0069】次に、図4に示すように、認識結果候補「東京都」が、候補確定窓23に表示されているときに、ジョイスティック12が押し込まれると、それが、入力音声「東京都品川区北品川」の始まりの部分「東京都」の認識結果として仮決定される。すると、認識部5において、辞書5aの標準パターンおよび構文情報が参照され、入力音声「東京都品川区北品川」の、既に仮決定された「東京都」に続く部分(入力音声の次の部分)に、辞書5aに記憶された標準パターンすべてが、その部分パターンとして、順次照合され、その結果、尤度の大きい上位N個の標準パターンが、入力音声の次の部分(「品川区」以下の部分)の認識結果候補とされる。

【0070】そして、この認識結果候補が、その尤度の大きい順に、例えば「豊島区」、「港区」、「品川区」、「千代田区」、「北区」、……であったとした場合、正しい認識結果としての「品川区」の尤度は最も大きくないが、即ちこの時点では誤認識であるが、これらの認識結果候補のうち、尤度が最も高かった標準パターンは、3文字の「豊島区」であるため、図6に示すように、ウインドウ24aの右横にオープンされたウインドウ24bに、3文字の認識結果候補「豊島区」、「品川区」、……だけが、尤度の大きい順に表示される。

【0071】その後、使用者によってジョイスティック12が操作され、認識結果候補として正しい「品川区」が表示された位置にカーソルが移動される。このとき、候補確定窓23には、既に仮決定された「東京都」に続いて、カーソルのある位置に表示されている「品川区」が表示される。そして、ジョイスティック12が押し込まれると、それが(「品川区」、入力音声「東京都品川区北品川」の、既に仮決定された部分「東京都」に続く部分の認識結果として仮決定される。

【0072】以下同様の処理が、入力音声の、既に仮決定された部分「東京都」および「品川区」に続く部分（「北品川」）に対して行われ、図7に示すように、ウィンドウ24bの右横にオープンされたウィンドウ24cに、認識結果候補「北品川」、「東品川」、・・・が表示される。その後、これらの認識結果候補のうちの、正しい認識結果としての「北品川」にカーソルが位置しているときに、ジョイスティック12が押し込まれると、それが（「北品川」）、入力音声「東京都品川区北品川」の、既に仮決定された部分「東京都」および「品川区」に続く部分の認識結果として仮決定される。

【0073】そして、入力音声に、既に仮決定された部分「東京都」、「品川区」、および「北品川」に続く部分がなくなり、図示せぬ確定ボタンが操作されると、候補確定窓23に表示された仮確定結果「東京都品川区北品川」が、入力音声の正しい認識結果とされる。

【0074】従って、認識部5において認識処理（本実施例においては、単語単位での認識処理）が行われるごとに、尤度の大きい順に出力される認識結果候補の数をN個とすると、従来の場合には、Nのべき乗で増加する認識結果候補を記憶しておかなければならなかったが、この音声認識装置によれば、認識処理ごとに、N個の認識結果候補のうちの、正しい認識結果を確定（仮決定）するようにしたので、Nの倍数で増加する認識結果候補を記憶すれば済み、よって、装置の小規模化を図ることができる。

【0075】以上のようにして認識結果「東京都品川区北品川」が得られると、システムコントローラ6において、CDROM9から地図情報が読み出され、「東京都品川区北品川」付近の地図が、適度な縮尺で表示部8に表示される。

【0076】以上、本発明の音声認識装置をナビゲーションシステムに適用した場合について説明したが、本発明はナビゲーションシステムの他、例えばオーディオ装置など、音声によって装置に動作を行わせる場合に適用することができる。

【0077】なお、本実施例においては、同一文字数の認識結果候補を、その尤度の大きい順に表示するようにしたが、例えば五十音順に表示したり、あるいはその両方を組み合わせて表示するようにすることができる。

【0078】さらに、本実施例では、表示部8に表示された認識結果候補の文字数を増減させるのにジョイスティック12を用いるようにしたが、例えばカーソルキーや、装置の筐体を引き張ったり押し潰したりする圧力を電気信号に変換する圧力トランスデューサなどを用いるようにしても良い。

【0079】また、本実施例においては、認識部5における音声認識を単語単位で行うようにしたが、例えばサブワードや音韻、文節、文などの単位で行うようにすることができる。さらに、サブワード単位や音韻単位など

で音声認識を行う場合、その認識結果を、例えば単語単位に変換し、表示部8に表示させるようにすることができる。

【0080】さらに、本実施例では、単語単位の認識結果候補を、その文字数で分類するようにしたが、例えば認識結果候補が、文節や文単位などで出力される場合には、認識結果候補を、それに含まれる単語数などで分類するようにすることができる。

【0081】また、本実施例においては、連続音声認識の場合について説明したが、本発明は、離散単語音声認識の場合についても適用することができる。

【0082】

【発明の効果】請求項1に記載の音声認識装置によれば、入力された音声が分析され、その分析結果に基づいて、音声認識されて、認識結果候補が出力される。そして、その認識結果候補が、所定の構成単位数に基づいて分類されて表示される。従って、認識結果候補の中から、正しい認識結果を容易に選択することができ、正しい認識結果を短時間で得ることができる。

【0083】請求項2に記載の音声認識装置によれば、表示手段に、選択手段により選択された所定の構成単位数の認識結果候補を表示させるので、認識結果候補の中から、正しい認識結果を、さらに容易に選択することができる。

【0084】請求項3に記載の音声認識装置によれば、表示手段に、分類手段により分類された認識結果候補のうちの最も尤度の大きいものと同一の所定の構成単位数の認識結果候補を表示させる。従って、より確からしい認識結果候補が表示されるので、正しい認識結果を、さらに短時間で得ることができる。

【図面の簡単な説明】

【図1】本発明の音声認識装置を適用したナビゲーションシステムの一実施例の構成を示すブロック図である。

【図2】図1の実施例の動作を説明するフローチャートである。

【図3】音声認識時の表示部8の表示画面を示す図である。

【図4】音声認識時の表示部8の表示画面を示す図である。

【図5】音声認識時の表示部8の表示画面を示す図である。

【図6】音声認識時の表示部8の表示画面を示す図である。

【図7】音声認識時の表示部8の表示画面を示す図である。

【図8】従来の音声認識装置の一例の構成を示すブロック図である。

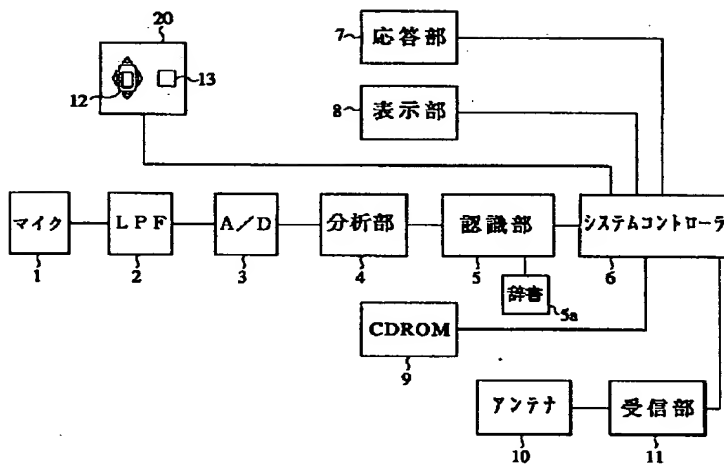
【符号の説明】

- 1 マイク
- 2 ローパスフィルタ（LPF）

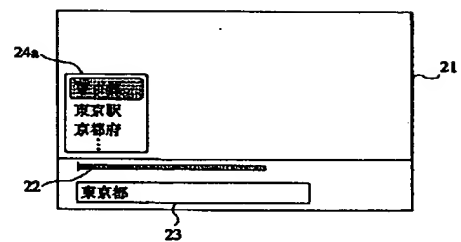
- 3 A/D変換器
- 4 分析部
- 5 認識部
- 5a 辞書
- 6 システムコントローラ
- 7 応答部
- 8 表示部
- 9 CDROM
- 10 アンテナ

- 11 受信部
- 12 ジョイスティック
- 13 ボタン
- 20 操作部
- 21 音声入力窓
- 22 インジケータ
- 23 候補確定窓
- 24a乃至24c ウィンドウ

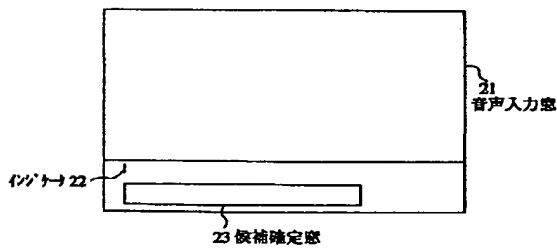
【図1】



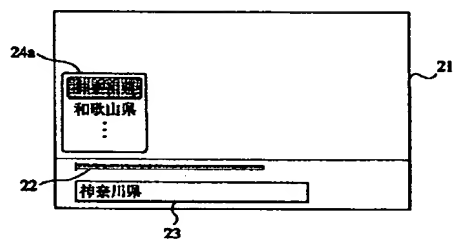
【図4】



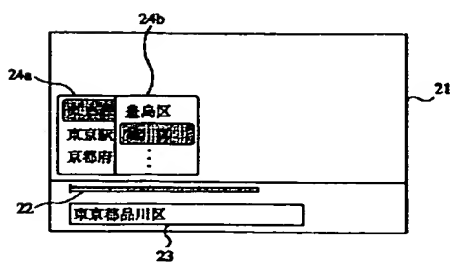
【図3】



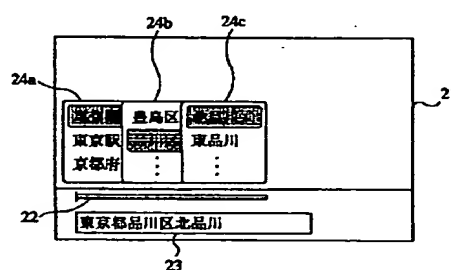
【図5】



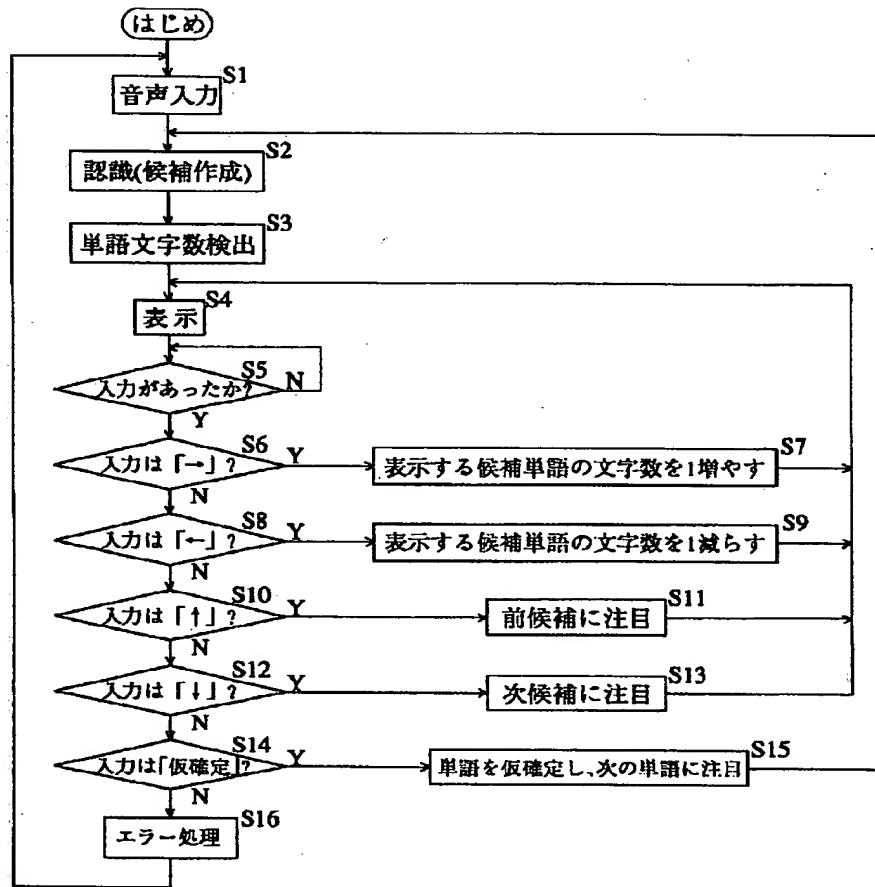
【図6】



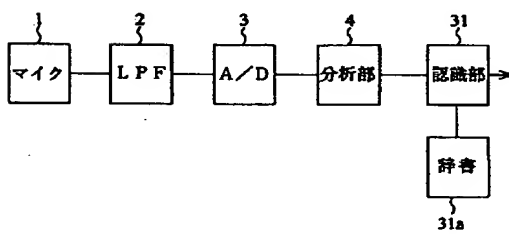
【図7】



【図2】



【図8】



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.